



# Big Data Survey Report

## SCORR Marketing and Applied Clinical Trials

**SCORR** + **APPLIED CLINICAL TRIALS**  
MARKETING



## Overview

“Big data” is used as a concept, catchphrase and buzzword in the drug development industry. In theory, it encourages analysis of very large datasets to help improve clinical trial outcomes and reduce costs. In practice, though, does it live up to the hype?

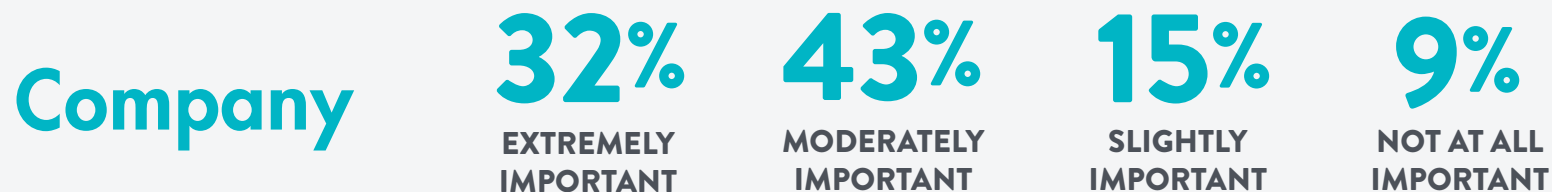
SCORR Marketing, in partnership with Applied Clinical Trials, recently conducted a survey to ascertain from industry professionals what big data comprises as well as its importance, how it is used today in the industry and how it may impact drug development services in the future.

The survey respondents are from contract research organizations (CROs), pharmaceutical companies, research sites, academic institutions, consultancies, biotech companies, medical device companies, service providers, technology companies and hospitals. Their departments or job titles include research and development (R&D), project management, clinical director, clinical research associate (CRA), corporate management, data management, strategy/planning, regulatory, IT/programming, business development/sales, medical affairs, quality assurance/quality control (QA/QC) and business development/sales. While a majority of respondents work for companies located in North America, there were also many respondents from European companies as well as some located in Asia, South America, Africa and Australia.

In this survey, we gathered information about:

- Big data’s importance and benefits
- What information comprises big data
- Who supplies big data to the industry
- Database network participation
- Concerns
- Stakeholder utilization and resistance
- Future effects

## How important is it that your company/drug development industry embrace the utilization of big data practices in clinical trials?



### What is your big data comprised of?

(Select all that apply.)

Utilization		Most Beneficial
77%	Medical records	48%
56%	Biomarker data	19%
43%	Registry data	13%
35%	Internet searches	8%
19%	Social networks	3%
15%	Financial records	2%
13%	Mobile sensors	1%
11%	Other	6%

### Which is most beneficial to you?

(Select one.)

## Which source(s) do you use to access big data?

(Select all that apply.)

Sources of Data Used	
Health care providers	57%
Internal creation/collection	56%
Third-party aggregators	38%
Insurance providers	16%
Other	7%

## Which one source do you use most often to access big data?

Single Most Used Source of Data	
Internal creation/collection	36%
Health care providers	32%
Third-party aggregators	24%
Other	5%
Insurance providers	3%


## Which of the following best or least describes what using big data in clinical trials means to you?

Best Describes	
Better understanding of study outcomes	20%
Improved protocol design	18%
Increased drug safety detection and monitoring	17%
More informed preclinical and translational research	15%

Least Describes	
Better patient retention	22%
Shorter studies	14%
Improved site selection	13%
More efficient patient recruitment	8%


## Rank the following stakeholders in terms of their utilization of big data practices:

(Where 1 represents most utilization, and 5 represents least utilization)

Top Users of Big Data	
	Average Ranking
 <b>PHARMACEUTICAL COMPANIES</b>	<b>1.96</b>
<b>CROS</b>	<b>2.62</b>
<b>THIRD-PARTY SERVICE PROVIDERS</b>	<b>3.36</b>
<b>CLINICAL SITES</b>	<b>3.37</b>
<b>LABORATORIES</b>	<b>3.48</b>

## Rank the following stakeholders in terms of their acceptance of using big data in clinical trials:

(Where 1 represents most acceptance, and 6 represents least acceptance)

Most Accepting of Big Data	
	Average Ranking
 <b>THIRD-PARTY SERVICE PROVIDERS</b>	<b>2.95</b>
<b>CROS</b>	<b>3.01</b>
<b>PHARMACEUTICAL COMPANIES</b>	<b>3.04</b>
<b>LABORATORIES</b>	<b>3.65</b>
<b>PATIENTS</b>	<b>4.29</b>
<b>CLINICAL SITES</b>	<b>4.47</b>

How will we use big data in five years?  
Which of these statements is most likely or least likely to happen?

	Most Likely	Least Likely
All new patient data will be in the form of electronic health records (EHRs).	32%	18%
Predictive modeling will be used to inform almost all preclinical and translational research.	23%	8%
Risk-based monitoring will be the new standard (instead of regular site visits).	22%	7%
More than half of clinical trial data will be shared.	10%	25%
Using mHealth wearables for data collection will be the norm.	7%	4%
Patient recruitment rates will be significantly better.	4%	29%
Big data won't significantly change the clinical trial landscape.	3%	9%

## Which of the following concerns have emanated from your company's use of big data?

(Select all that apply.)

**46%**

DATA INTEGRITY CONCERNS

**44%**

DATA SECURITY CONCERNS

**41%**

PATIENTS' RIGHT TO PRIVACY

**40%**

LEGAL CONCERNS

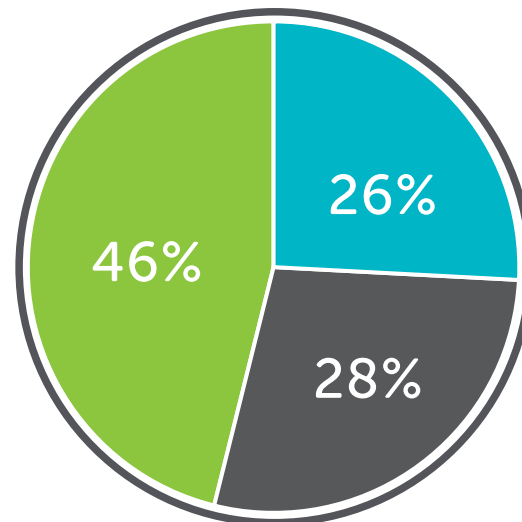
**35%**

INTELLECTUAL PROPERTY CONCERNS

**22%**

NO CONCERNS

## Does your company participate in a database network (to share and to gain access to clinical trial data)?



Yes

No

Unsure

## Three Main Points:

### 1. While respondents say their companies should embrace big data practices, they believe it is even more imperative for the industry as a whole to do so.

Three-fourths of those surveyed (75%) believe that it is either extremely important (32%) or moderately important (43%) for their own companies to adopt big data practices in clinical trials. Yet even more say the industry must embrace big data: Nine in 10 respondents (90%) either believe that this is extremely important (58%) or moderately important (32%).

These numbers suggest that many who took the survey want the industry to use big data more regularly even if they don't see a need or a specific application for their own companies.

When we compare the opinions of those who strongly believe the industry should adopt big data to the views of those who say the industry's adoption is of less importance, some interesting differences come to light. Those respondents who strongly favor industry usage of big data:

- Are more likely to use internally created or collected information as a source of big data (62%) than those who consider industry adoption of big data to be less important (48%). On the other hand, they are less likely to use health care as a data source (52%) than those on the other end of the spectrum (65%).

- Have greater data security concerns (48%) but fewer data integrity qualms (43%) than do the people who consider the industry's use of big data as only slightly or not at all important (40% concerned about data security; 51% uneasy about data integrity).
- Are much more inclined to believe that risk-based monitoring will be the new standard (instead of regular site visits) in five years (26%) than are those who place less importance on industry adoption of big data (15%).

### 2. "Big data" means different things to different people.

Survey respondents were asked to identify which one of nine options best describes what using big data could mean. The four options most selected are:

1. Better understanding of study outcomes
2. Improved protocol design
3. Increased drug safety detection and monitoring
4. More informed preclinical and translational research

A plurality of support for each of these options is led by different groups. That some answer choices are correlated with specific groups is not surprising. Each of these groups understandably highlights potential benefits that would affect what they do.

- Better understanding of study outcomes: 38 percent of contract research organizations (CROs) selected this answer, while 20 percent of all respondents did.



## Three Main Points (cont):

- Improved protocol design: 43 percent of biotechs chose this option, which was picked by only 18 percent of all respondents overall.
- Increased drug safety detection and monitoring: Indicated by 35 percent of pharmaceutical company respondents but by only 17 percent of all those who took the survey.
- More informed preclinical and translational research: Chosen by 35 percent of academic institutions and 15 percent of respondents overall.

Concerns identified over big data usage are also dependent on the respondent's company or position. For example, those who work in data management positions are especially inclined to voice concerns over data integrity (79%). Those from academic institutions are much more likely to be uneasy about data security (71%).

### 3. Biotech companies appear to be not only less supportive of but also less concerned about big data than any other group.

Only a little more than half (54%) of biotech respondents consider it extremely or moderately important for their company to embrace big data practices in clinical trials. This is noticeably lower than other groups: pharmaceutical companies (91%), research sites (80%), academic institutions (77%) and CROs (76%).

Similarly, biotech companies are more apathetic about the future impact of big data. A quarter of biotech representatives (25%) indicated that "big data won't significantly change the clinical trial landscape." Fewer than one in 10 members of academic institutions, CROs, pharmaceutical companies and research sites selected this answer (from among seven options).

Other responses from biotech respondents also pointed to this group's relative indifference toward big data:

- On average, respondents named 2.69 types of information that comprise big data. Biotech representatives named only 2.32. (In contrast, those from pharma named 3.26 on average.)
- Overall, the mean number of sources of big data named by those surveyed is 1.74. The mean for biotechs is 1.25 (compared to 2.00 for pharma).
- The average number of identified concerns overall is 2.29, which is the same average as for pharma alone, while for biotechs, it is 1.50.

Finally, biotech companies were much more likely to answer "no opinion" or "no concerns" when asked about potential benefits or concerns. More than one-third (36%) of biotechs had no opinion when asked to select from a number of benefits, and one-third (33%) expressed no apprehension when asked to choose from a variety of potential concerns.

## Other Key Takeaways:

- Medical records are most associated with big data for all survey sub-groups. Biomarker data rates a distant second.
- More than one-half of all respondents use health care providers as a source of data or internally create or collect data to access big data.
- Respondents were least likely to say that using big data means better patient retention or improved site selection.
- About a quarter of respondent companies participate in a database network to share clinical trial data with other companies in order to gain access to a larger set of data. Slightly less than half do not, while the remaining respondents were uncertain.
- Pharmaceutical companies are, by far, the stakeholder most identified as utilizing big data practices.
- Clinical sites and patients are identified as the groups most resistant to the use of big data in clinical trials. Biotech and those in data management positions were most likely to choose clinical sites as the group most hesitant, while CROs and project managers were most likely to pick patients.
- Just one in 25 respondents believe that the application of big data will significantly improve patient recruitment rates over the next five years. This surprising lack of support is universal; for each identified organization type, job position or company location, fewer than one in 10 respondents believe that big data would significantly improve patient recruitment.